DOCUMENT RESUME

ED 175 290 FL 010 723

AUTHOR Larsen-Freeman, Diane; Strom, Virginia
TITLE The Construction of a Second Language Acquisition Index of Development.
PUB DATE 77
NOTE 13p.; Paper presented at the Los Angeles Second Language Research Forum (1st, Los Angeles, California, February 11-13, 1977)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Rating; College Students; Communicative Competence (Languages); *Composition (Literary); *English (Second Language); *Error Analysis (Language); Grammar; Language Development; Language Proficiency; Language Research; Rating Scales; *Second Language Learning; Sentences; *Writing Skills

ABSTRACT

Compositions written by 48 university students of English as a Second Language (ESL) were examined as a step in the development of an index for proficiency in a second language. A feature analysis of the compositions revealed the following tendencies: (1) syntactic sophistication and tense usage improved as proficiency increased, (2) errors in morphology decreased between the "fair" and "excellent" levels, (3) prepositions were a consistent problem throughout all five levels, and (4) the number of errors in articles first increased and then decreased as the quality of the compositions improved. An analysis based on the relationship between features and length of composition produced the following results: average number of words per composition increased with proficiency, and the variance in number of words per sentence from one level to another, while striking, did not show a clear trend across levels of proficiency. Finally, the average number of error-free sentences increased significantly as proficiency level increased. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
L.A. 2nd Lang. Res. Forum
Carol Henning, Ed.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED175290

FL010723

The Construction of a Second Language Acquisition Index of Development[1]

Diane Larsen-Freeman and Virginia Strom[2]
University of California, Los Angeles

Introduction

For some time second language acquisition researchers have acknowledged the need for an index of development by which a learner's proficiency in the second language could be guaged. First language acquisition researchers have been able to employ a measure of the linguistic development of their subjects by calculating their mean length of utterance in number of morphemes. Although such a measure might not be totally satisfactory, even this measure is not applicable in second language acquisition research where the learner is more cognitively mature and therefore capable of producing utterances which are more than a few morphemes in length shortly after he has had initial contact with the target language.

While we would be the first to admit that individual learner variablilty is evident in second language learning, we are also convinced there exist general developmental progressions or trends to which learners conform. Furthermore, we believe the delineation of stages in this progression would allow us to address questions which we are currently forced to ignore. Hakuta (1975), for example, conducted a longitudinal study of the acquisition of English by Uguisu, a five-year old native Japanese speaker. One of the issues he chose to examine was language transfer as a factor in the second language learning process. He was specifically interested in the structural avoidance phenomenon observed by Schachter (1975).

In order to study this, he decided to compare Uguisu's use of relative clauses with that of Marta, a five-year old Spanish-speaker learning English as a second language who was then a subject of a study by Cancino, Rosansky and Schumann (1975). He hypothesized that since in Spanish like English, and unlike in Japanese, the relative clause occupies the position to the right of head noun, Marta, the Spanish-speaker, would produce more relative clauses than Uguisu, the Japanese-speaker.

He was unable to test his hypothesis, however, since he had no reliable means of determining whether Uguisu and Marta were at a

---

[1]Paper presented at the First Annual Second Language Research Forum, UCLA, Los Angeles, CA. February, 1977.

[2]The authors are grateful for the assistance of Nan Wonderly and Earl Rand in the preparation of this paper.

similar stage of development in English. Marta might have been producing more relative clauses simply because she had a better command of English than Uguisu, not because she spoke a first language more similar to English than Japanese. Had he been comparing the number of relative clauses used by two children learning English as their first language, he could have been reasonably safe in assuming that they were at a similar stage of language development since they were both five years of age. Unfortunately, we in second language acquisition research cannot even rely upon the independent yardstick of chronological age by which to compare subjects.

Taylor (1975) and Larsen-Freeman (1975,1976) have both demonstrated that the strategies of the learner, or at least their influences on the learning process, change as the learner's proficiency in the target languages increases.

Taylor, for example, administered a translation test to twenty Spanish-speaking subjects of English as a second language. The subjects were at two levels of proficiency--"beginning" and "intermediate." An error analysis made of their translations revealed that the errors committed by the elementary and intermediate level students were not qualitatively different--however their relative influence depended on the level of proficiency of the learner. For instance, the degree to which the subjects relied on the strategies of overgeneralization and transfer were quantitively different--the elementary subjects' reliance on the transfer strategy was found to be significantly higher than that of the intermediate subjects. Thus, when making any claims about the learner, researchers must specify at which stage of development their subjects are operating. We tend to get embroiled in controversies because we overlook this. We argue about the influence of a certain factor on the language learning process while overlooking the probability that the influence of that factor might fluctuate depending upon the learner's proficiency level.

Those of us involved in second language acquisition research usually do report the target language proficiency of our subjects, but we use very subjective labels. We identify learners as being beginning, intermediate or advanced with respect to the target language. Such a classification schema is much too vague and relative to convey to others the level of competence of a particular group of learners. While we acknowledge the danger inherent in attempting to divide a process such as language learning into discrete units such as stages, we believe the risk a necessary one which will enable researchers to explicitly identify the population of learners with whom they are dealing.

Procedures

The remainder of this paper will be a progress report on our attempt to construct a second language index of development. We have chosen to examine the acquisition of English as a second language, but it was our hope that we might find some measure of language proficiency that could be applied universally.

We started initially with written compositions, because they were accessible and written data seemed easier to work with than spoken. However if we are successful in this area, we would eventually like to expand our efforts and attempt to characterize stages of development in communicative speech.

Compositions written by forty-eight non-native speakers of English were examined. The authors were undergraduate and graduate students at UCLA. There were 28 males and 20 females in our sample. The compositions were written for the UCLA English as a Second Language Placement Examination. Our only criterion for choosing these particular compositions was that there would be no more than two written by speakers of the same language. We wanted a diversity of language backgrounds in order to allow us to construct a global index of development. The compositions were written under examination conditions. The time spent on writing them was not restricted by the examination itself, although a minimum length was assigned in each case.

On an impressionistic basis two researchers independently assigned each composition to one of five levels of proficiency: poor, fair, average, good, excellent. The inter-researcher reliability was found to be statistically significant ($r=.9713$).

Because the agreement between researchers was so high we felt justified in taking the next step--to try and identify the features that made each group of compositions unique.

Thus, the compositions were analyzed in order to identify features that made each level distinct from the others. Features of each composition were considered--both those that resulted in lack of grammaticality (i.e., errors) and those that made the composition appear more native-like. One set of features considered was writing mechanics. Punctuation, capitalization, and spelling were included in this category. Organization, clarity and syntactic sophistication were also assessed. For example, one of our subjects wrote a composition containing a sentence which read: "I was surprised by the openness with which he described himself." We recorded sentences of this caliber [+ syntactic sophistication] each time they were encountered.

Lexical choice was another characteristic considered. One of our subjects wrote: "The employers wanted to press down the wages." What he probably meant to say was "hold down." Such instances were recorded as [- lexical choice].

We scrutinized each composition for the author's ability to write grammatically. More specifically we looked at morphology, syntax, prepositions, tenses, aspect, articles, subject/verb agreement, case, and negation. Sentence construction violations such as fragments and run-ons were noted. The content of the composition was also evaluated as to relevance to the topic, degree of interest, and sophistication of treatment.

Although we were able to isolate features and errors which made the composition more or less native-like, we realized an index of development based solely on these would have certain drawbacks.

First: How would errors or positive features such as those just listed be weighted? Would omission of an article from an obligatory context, for example, be considered more serious than the use of an inappropriate lexical item? Being forced to make somewhat arbitrary decisions such as these disturbed us.

Second: If we were to establish an index of development based on the decreasing number of errors committed with regards to one or any group of linguistic structures, we would have to assume linear development in language acquisition. Errors in articles, for example, would have to decrease regularly as proficiency in English increased. Henning (1975), Hakuta (1975) and other researchers have shown this is not the case in the acquisition of grammatical structures. If one were to plot the acquisition of a structure by an individual learner over time, one would note that the line is full of peaks and valleys. Our data, too, provide evidence that learning does not progress linearly. For example, we discovered the subjects writing compositions we had evaluated as "poor" committed fewer errors in article usage than those subjects who wrote compositions we had evaluated as "fair." It was only at the level of compositions determined to be "good" and "excellent" that the number of errors in article usage declined to the level they were at for the writers of the "poor" compositions. If we used errors in article usage alone as our index of development, we would have no way of discriminating between the "poor" and "good" levels of writing.

Neuman (1976), too, encountered this problem when she attempted to characterize the problems of learners studying English as a second language at the intermediate level at UCLA. After a thorough error analysis of compositions, she concluded errors cannot be used to distinguish the intermediate level from the beginning and advanced levels. The only area that distinguished among the groups was problems with word order. However, even this would not serve as a basis for an index of development since although the greatest number of errors were made by the beginning learners, the advanced learners made somewhat more errors than those at the intermediate level.

Third: If we based our index on the absence of commission of errors in particular structures, we might very well be discriminating against one particular native language group who might experience more interference with that structure than other language groups.

Thus, in an attempt to find something which could be used as an index of development which would be exempt from these constraints, we turned to more objective features--measures of length.

The number of words and the number of sentences were counted in each subject's composition. We also counted the total number of T-units

per composition. A T-unit is a "minimal terminable unit, ...minimal as to length, and each...gramatically capable of being terminated with a capital letter and a period." (Kellogg W. Hunt, Grammatical Structures Written at Three Grade Levels, 1965) Within a single sentence, therefore, more than one T-unit may occur. For example, a sentence such as the following would contain two T-units: "I first met this very special man when I was only about a year old, and my earliest memories of him are vague."

Next, we determined the average word length per T-unit per subject. The reason the length of T-units was thought to be important was Hunt's finding that as children learning English as their first language mature, they tend in their writing to reduce main clauses to structures such as words and phrases which are incorporated within other main clauses. This tendency, Hunt reports, increases almost in a straight line. That is, the number of sentences they write diminishes, but the length of the T-units increases due to embedding and other processes which demote main clauses to a subordinate status.

The number of error-free T-units was counted as well. We were prompted to do this after recalling a study conducted by Scott and Tucker (1974). They considered the number of error-free T-units along with many other factors in their error analysis of compositions written by Arab students learning English as a second language.

Although Scott and Tucker judged a T-unit to be error-free if correct in all syntax and function words, we only counted it error-free if the T-unit was perfect in all respects, including spelling and punctuation.

We had not thought it necessary up to this point to be concerned about the differing lengths of the compositions. But with the number of error-free T-units, we thought it important to control for length. Thus, we did our analysis on only thirty-seven of the forty-eight compositions--only those whose authors were instructed to write 200 words.

We created charts for each proficiency level displaying the errors and features of each of the five groups. A one-way analysis of variance was used to test for significant differences between groups for the average number of words per T-unit per subject. A one-way analysis of variance was also computed for the number of error-free T-units among the 37 subjects instructed to write 200-word compositions.

Results

From the first part of our study, the analysis of errors and positive features, it is difficult to draw conclusions.

We recorded every conceivable error and every positive feature, but since the number of obligatory contexts for each of these varied with each subject, obviously the number of errors in each category

would fluctuate from subject to subject.

Nevertheless, from a preliminary inspection, we were able to discern some trends among all the things we investigated. There was nothing startling. Spelling improved as proficiency increased. The number of errors went down considerably, even while lexical choice improved (that is, subjects were using more sophisticated, less common words and spelling them correctly). Learners are acquiring a system as opposed to learning to spell individual items.

Syntactic sophistication improved--from a basic NP + VP pattern at the poor level to sentences with multiple embeddings of relative clauses, gerunds, infinitives, etc. at the excellent level.

Tense usage (i.e., appropriateness in form and function) improved as proficiency increased.

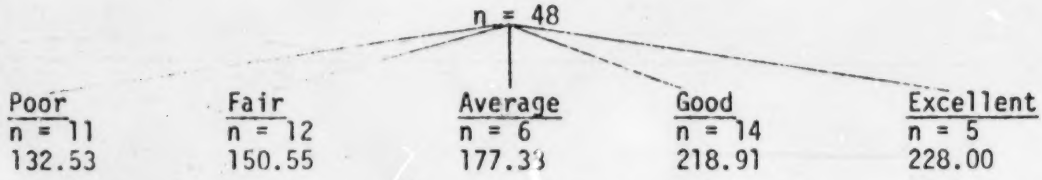
Errors in morphology decreased slightly from those compositions judged to be "fair" to those evaluated as "excellent."

Prepositions were a consistent problem throughout all five levels. There was no noticeable difference in the commission of errors among the five levels.

As we reported earlier, there was an interesting trend with errors in articles. The "poor" compositions had a certain number of errors; the "fair" ones had more errors than the "poor" but then the number of errors decreased again, i.e., there was no uniform linear progression.

Obviously, we will have to examine the errors and positive features more thoroughly--but for now, let us turn to the results of the quantitative analysis.

1. Average number of words per composition



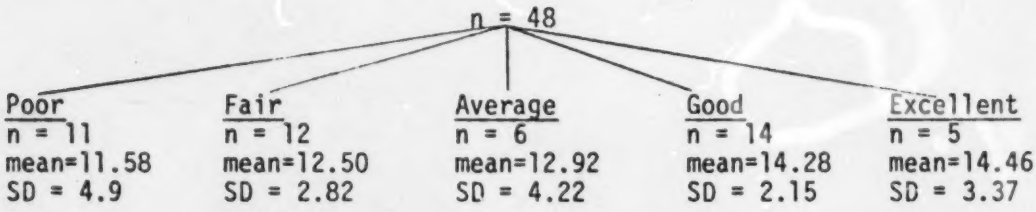
This is an interesting finding in and of itself. The subjects were not constrained in time and yet there is a noticeable correlation between the number of words and our evaluation of their proficiency.

Neuman (1976) found this, too. Her beginning subjects wrote an average of 129.85 words per composition, her intermediate subjects wrote 214.75, and her advanced subjects 286.75. (n = 20)

Maybe sheer length of a composition is distinctive. The longer length might be attributed to a better command of syntax or vocabulary by a subject or simply that the subjects with greater confidence in their

English ability are willing to be more expressive and take greater risks of committing errors.

2. Average number of words per T-unit



The one-way analysis of variance revealed the probability the F-value was exceeded as .32. The results were not significant--the groups were not statistically different from one another.

However, Figure 1 shows that the mean length of T-units does seem to increase fairly steadily, although it is statistically insignificant.

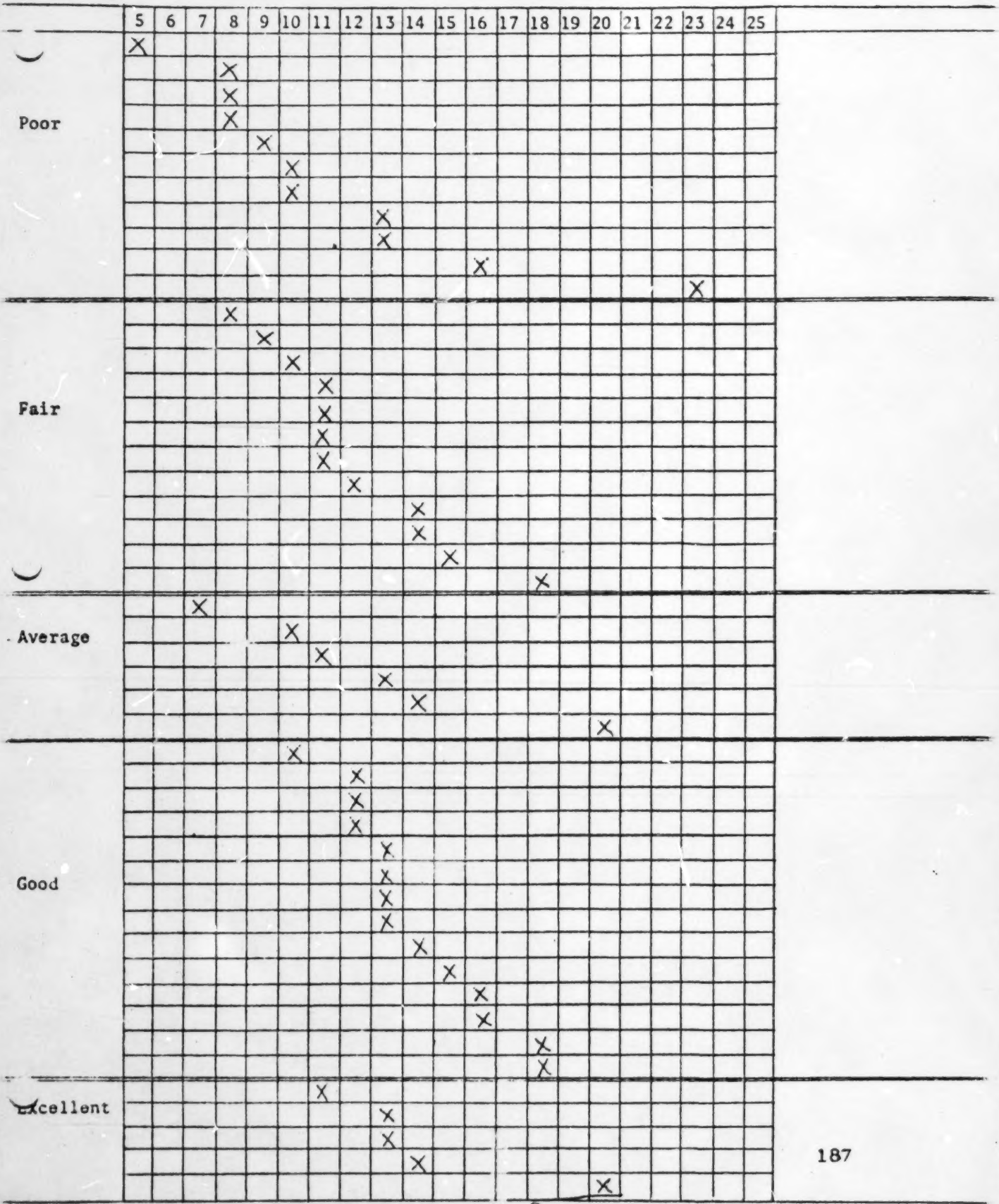
If the extremes were eliminated from our sample, the means would probably be adjusted downward. Furthermore, it is interesting to note that the fairly obvious modes for the "poor" (mode = 8), "fair" (mode = 11), and "good" (mode = 13) compositions do increase regularly in the desired direction. Perhaps with a larger sample, this pattern would be more impressive.

The big factor affecting significance in T-unit length was probably the size of the standard deviations.

When the compositions were considered for the average number of words per T-unit, there were a few compositions that deviated from the pattern for each level. That is, there were two "poor" compositions with a large number of average words per T-unit and others that were "good" or "excellent" with a small number of words per T-unit. Some of the variance is no doubt due to the fact that is was difficult to identify T-units in some of the "poor" compositions because of their ungrammaticality. For example, when subjects attempted to use relative clauses without having mastered how to do so, it was difficult to determine where one T-unit stopped and another began. This might artificially inflate the T-unit length and would account for a low rating with a large number of words per T-unit.

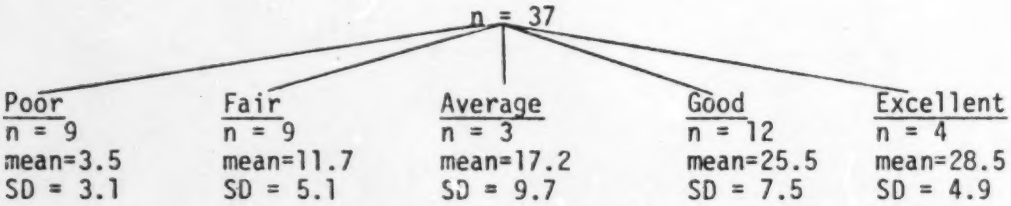
A couple of "excellent" compositions contain short T-units. One of these has excellent grammar and spelling but not as much conjoining or relativizing as there could be. The other used a great deal of dialogue, which makes the composition very interesting but does not allow for long T-units.

Figure 1
Average Number of Words Per T-Unit
(n = 48)



187

3. Average number of error-free T-units (200-word assignment)



The one-way analysis of variance revealed that the probability of the F-value being exceeded was .000--highly significant--among the 37 writers who had received identical instructions, i.e., to write 200-word compositions.

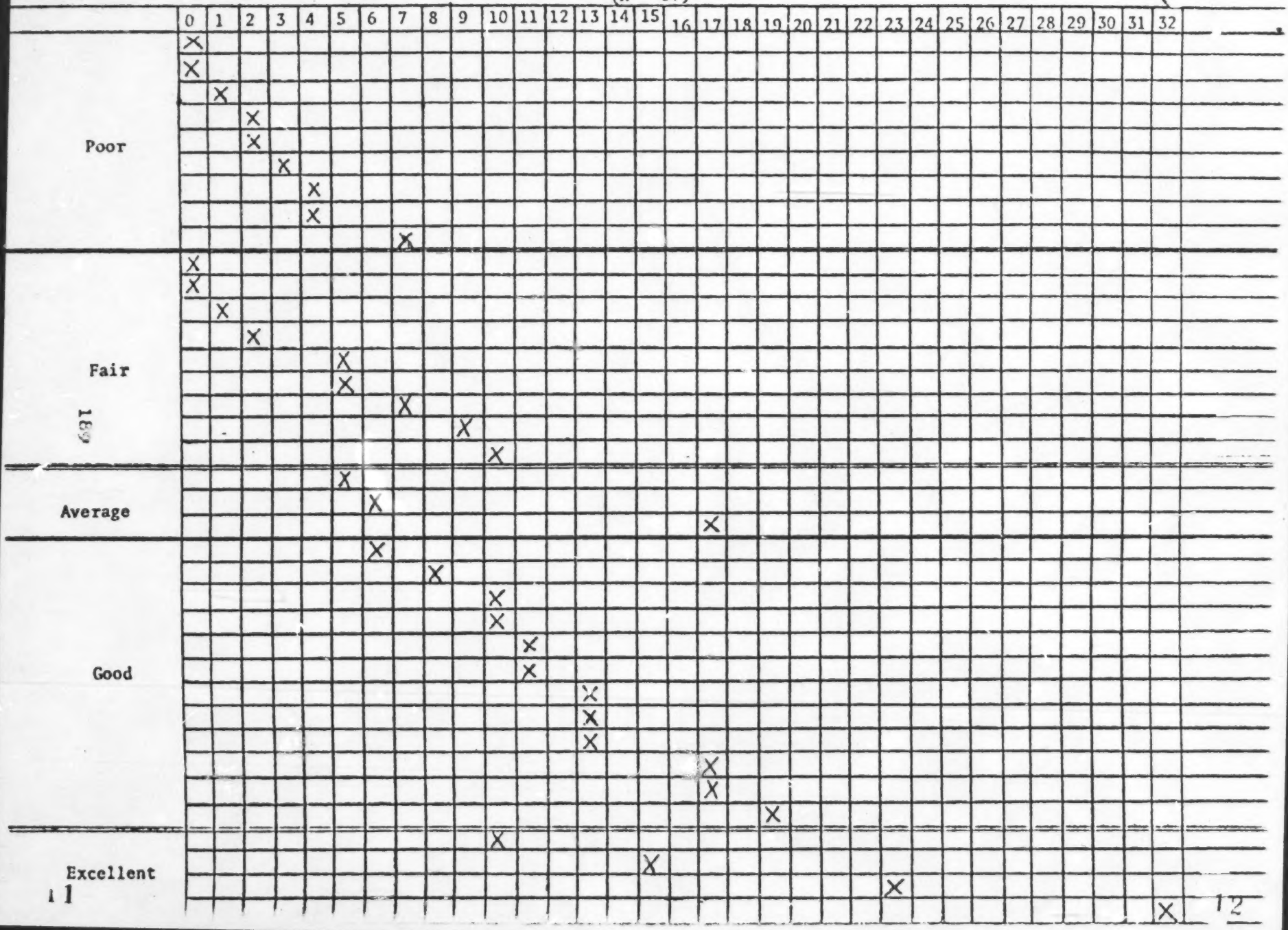
An approximation of the straight line progression we sought can perhaps be appreciated from the next chart (see figure 2). We notice there is still a good deal of variance but there is also an unmistakeable trend toward a greater number of error-free T-units as proficiency level increases.

Conclusions

Error analysis and positive feature analysis do tend to reveal some patterns in the manipulation of structures and features as proficiency in composition writing increases. They warrant more attention on our part. We are not optimistic, however, that the reduction of errors in any particular structure or group of structures will be the answer to our quest for an index of development. Analyses of this sort may provide evidence which will serve as further indication of proficiency supplementing something such as length of T-unit or number of error-free T-units. At this point the measures of length and error-free T-units seem to be much more viable contenders on which to base an index of development since they are easily quantified, linear progression does seem possible and they would seem to be impervious to differences in language backgrounds.

Our next step will be to correlate the best possible indicator of proficiency with available ESL placement test scores of the university students to see if there is a significant positive correlation. Then we plan to increase the number of subjects and do further research on length measures--both with subjects within the university as we have done, and also with subjects from other age groups and backgrounds. With other populations we may very well find that the number of error-free T-units is too low per composition to be an effective measure of individual development. Finally, if the establishment of an index of development is possible for written compositions, we would like to next turn and attempt to do the same with communicative speech. The end product will hopefully be a second language acquisition index of development by which we can all adhere.

(200-word assignment)
(n = 57)



189

11

12

Bibliography

Hakuta, K. 1975. "Becoming Bilingual at Age Five: The Story of Uguisu." Unpublished thesis, Harvard College.

Henning, G. 1976. "A Developmental Analysis of Errors of Adult Iranian Students of English as a Foreign Language." Paper presented at the UCLA-USC Second Language Acquisition Forum, Los Angeles, January 13, 1976.

Hunt, K. 1965. Grammatical Structures Written at Three Grade Levels (Research Report No. 3). Champaign, Ill.: National Council of Teachers of English.

Larsen-Freeman, D. 1975. "The Acquisition of Grammatical Morphemes by Adult Learners of English as a Second Language." Unpublished Ph.D. Dissertation, The University of Michigan.

Larsen-Freeman, D. 1976. "Evidence of the Need for a Second Language Acquisition Index of Development," to appear in Principles of Second Language Learning and Teaching, William Ritchie, editor. New York: Academic Press, Inc.

Neuman, R. 1977. "An Attempt to Define Through Error Analysis the Intermediate ESL Level at UCLA." Unpublished thesis, UCLA.

Schachter, J. 1974. "An Error in Error Analysis," Language Learning, 24, 205-214.

Scott, M. and Tucker, G.R. 1974. "Error Analysis and English-Language Strategies of Arab Students," Language Learning, 24, 69-97.

Taylor, B. 1975. "The Use of Overgeneralization and Transfer Learning Strategies by Elementary and Intermediate Students in ESL," Language Learning, 25, 73-107.